

Numerical study of tree-level improved lattice gradient flows in pure Yang-Mills theory

Norihiko Kamata* and Shoichi Sasaki†

Department of Physics, Tohoku University,
Sendai 980-8578, Japan

(Dated: September 30, 2016)

We study several types of tree-level improvement in the Yang-Mills gradient flow method in order to reduce the lattice discretization errors in line with a reference [Fodor *et al.*, arXiv:1406.0827]. The tree-level $\mathcal{O}(a^2)$ improvement can be achieved in a simple manner, where an appropriate weighted average is computed between the plaquette and clover-leaf definitions of the action density $\langle E(t) \rangle$ measured at every flow time t . We further develop the idea of achieving the tree-level $\mathcal{O}(a^4)$ improvement within a usage of actions consisting of the 1×1 plaquette and 1×2 planar loop for both the flow and gauge actions. For testing our proposal, we present numerical results of $\langle E(t) \rangle$ obtained on gauge configurations generated with the Wilson and Iwasaki gauge actions at three lattice spacings ($a \approx 0.1, 0.07$ and 0.05 fm). Our results show that tree-level improved flows significantly eliminate the discretization corrections on $t^2 \langle E(t) \rangle$ in the relatively small- t regime for up to $t \gtrsim a^2$. To demonstrate the feasibility of our tree-level improvement proposal, we also study the scaling behavior of the dimensionless combinations of the $\Lambda_{\overline{\text{MS}}}$ parameter and the new reference scale t_X , which is defined through $t_X^2 \langle E(t_X) \rangle = X$ for the smaller X , *e.g.* $X = 0.15$. It is found that $\sqrt{t_{0.15} \Lambda_{\overline{\text{MS}}}}$ shows a nearly perfect scaling behavior as a function of a^2 regardless of types of the gauge action and the flow, after tree-level improvement is achieved up to $\mathcal{O}(a^4)$. Further detailed study of the scaling behavior exposes the presence of the remnant $\mathcal{O}(g^{2n} a^2)$ corrections, which are beyond the tree-level. Although our proposal is not enough to eliminate the $\mathcal{O}(a^2)$ effects exactly, we show that the $\mathcal{O}(g^{2n} a^2)$ corrections can be well under control even by the simplest tree-level $\mathcal{O}(a^2)$ improved flow.

I. INTRODUCTION

Recently, the Yang-Mills gradient flow method [1] has continued to develop remarkably. Indeed, this method is extremely useful for setting a reference scale [1–3], computing the non-perturbative running of the coupling constant [4], definition of energy momentum tensor (EMT) on the lattice [5, 6], calculation of thermodynamics quantities [7] and so on [8]. These applications are based on measuring the expectation value of the action density $E(t, x)$. However, in calculation of $\langle E(t) \rangle$, there is still room for improvement with respect to the lattice gradient flow, where some lattice artifacts are found to be non-negligible [9, 10]. Therefore, it is important to understand how to practically reduce the effects of the lattice artifacts due to the finite lattice spacing a .

At tree level in the gauge coupling, lattice discretization effects on the expectation value $\langle E(t) \rangle$ had been already studied in the recent years [9, 10]. According to their results, tree-level discretization errors become large in the small flow-time “ t ” regime as inverse powers of t/a^2 . This tendency is problematic when we construct the lattice EMT operator using the Yang-Mills gradient flow and then calculate thermodynamics quantities such as trace anomaly and entropy density following Suzuki’s proposal [5]. The idea of Suzuki method is based on the

fact that flowed observables, which live on $4 + 1$ dimensional space, can be expanded by a series of the expectation values of the ordinary four-dimensional operator in powers of the flow time t , (so-called “small- t expansion”) [8]. Therefore, it is important to control tree-level lattice discretization errors on the action density $E(t, x)$, which is a key ingredient to evaluate the trace anomaly term in the lowest-order formula of the new EMT construction [5].

In this context, we would like to know what is an optimal combination of choices of the flow, the gauge action and the action density in line with the tree-level improvement for the lattice gradient flow [9]. A simple idea of achieving $\mathcal{O}(a^2)$ improvement was considered by FlowQCD Collaboration [11]. The appropriate weighted average of the values $t^2 \langle E(t) \rangle$, which are obtained by the plaquette and clover lattice versions of $E(t, x)$, can easily cancel their $\mathcal{O}(a^2)$ corrections. The weight combination was determined at tree-level in Refs. [9] and [10]. We have developed this idea to achieve tree-level $\mathcal{O}(a^4)$ improvement using the flow action consisting of both the plaquette and rectangle terms in Ref. [12].

In our previous work, it is found that although our tree-level improvement program, which is only valid for $t \gtrsim a^2$, significantly eliminate the discretization corrections in the almost whole range of t , some discretization uncertainties still remain in the large- t regime [12]. In this paper, we include additional numerical simulations with the renormalization group (RG) improved gauge action [13] and then extend our research scope to address the feasibility of our improvement program and also fully

*Electronic address: kamata@nucl.phys.tohoku.ac.jp

†Electronic address: ssasaki@nucl.phys.tohoku.ac.jp

understand the origin of the remnant discretization errors found in our previous work.

This paper is organized as follows. In Sec. II. after a brief introduction of the Yang-Mills gradient flow and its tree-level discretization effects, we describe our proposal where the tree-level improvement is achieved up to and including $\mathcal{O}(a^4)$ in a simple manner based on Ref. [9]. In Sec. III we show numerical results obtained from the pure Yang-Mills lattice simulations using two different gauge actions: the standard Wilson gauge action and the RG improved Iwasaki gauge action [13]. Section IV gives the details of the scaling study of the new reference scale defined in the Yang-Mills gradient flow method. We then discuss the remnant $\mathcal{O}(g^{2n}a^2)$ corrections, which are beyond the tree-level. Finally we summarize our study in Sec. V.

II. THEORETICAL FRAMEWORK

A. The Yang-Mills gradient flow and its tree-level discretization effects

Let us briefly review the Yang-Mills gradient flow and its tree-level discretization corrections. The Yang-Mills gradient flow is a kind of diffusion equation where the gauge fields $A_\mu(t, x)$ evolve smoothly as a function of fictitious time t . It is expressed by the following equation,

$$\frac{dA_\mu(t, x)}{dt} = -\frac{\delta S_{YM}[A]}{\delta A_\mu(t, x)}, \quad (1)$$

where $S_{YM}[A]$ denotes the pure Yang-Mills action defined in terms of the flowed gauge fields $A_\mu(t, x)$. The initial condition of the flow equation at $t = 0$, $A_\mu(0, x)$, is supposed to correspond to the gauge fields of the 4-dimensional pure Yang-Mills theory. Through above flow equation, the gauge fields can be smeared out over the sphere with a radius roughly equal to $\sqrt{8t}$ in the ordinary 4-dimensional space-time. One of the major benefits of the Yang-Mills gradient flow, is that correlation functions of the flowed gauge fields $A_\mu(t, x)$ have no ultraviolet (UV) divergence for a positive flow time ($t > 0$) under standard renormalization [1, 14].

To see this remarkable feature, let us consider a specific quantity, like the action density $E(t, x)$ that is defined by $E(t, x) = \frac{1}{2}\text{Tr}\{G_{\mu\nu}(t, x)G_{\mu\nu}(t, x)\}$. Here, the field strength of the the flowed gauge fields is given by

$G_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ ($\mu, \nu = 1, 2, 3, 4$) in the continuum expression. Taking the smaller value of t implies the consideration of high-energy behavior of the theory. Therefore, the vacuum expectation of $E(t, x)$ in the small- t regime, where the gauge coupling becomes small, can be evaluated in perturbation theory. In Lüscher's original paper [1], $\langle E \rangle$ was given at the next-to-leading-order (NLO) in powers of the renormalized coupling in the $\overline{\text{MS}}$ scheme, while its next-to-NLO (NNLO) correction has recently been evaluated by Harlander and Neumann [15].

The dimensionless combination $t^2\langle E(t) \rangle$ is expressed in terms of the $\overline{\text{MS}}$ running coupling g at a scale of $q = 1/\sqrt{8t}$ for the pure $SU(3)$ Yang-Mills theory:

$$t^2\langle E(t) \rangle = \frac{3g^2(q)}{16\pi^2} \left[1 + \frac{k_1}{4\pi} g^2(q) + \frac{k_2}{(4\pi)^2} g^4(q) + \mathcal{O}(g^6(q)) \right], \quad (2)$$

where the NLO coefficient k_1 was obtained analytically as $k_1 = 1.0978$ [1], while the NNLO coefficient k_2 has been evaluated with the aid of numerical integration as $k_2 = -0.982$ [15]. Unlike the ordinary 4-dimensional gauge theory, Eq. (2) has no term proportional to $1/(4-d)$, which is divergent in the limit of $d \rightarrow 4$, at this order. This UV finiteness has been proved not only for the above particular quantity at this given order, but also for any correlation functions composed of the flowed gauge fields at all orders of the gauge coupling [14].

The lattice version of $t^2\langle E(t) \rangle$ obtained in numerical simulations shows a monotonically increasing behavior as a function of the flow time t and also good scaling behavior with consistent values of the continuum perturbative calculation (2) that suggests the presence of the proper continuum limit [1]. The observed properties of $\langle E(t) \rangle$ offer a new reference scale t_X , is given by the solution of the following equation

$$t^2\langle E(t) \rangle|_{t=t_X} = X, \quad (3)$$

where $X = 0.3$ was adopted in Ref. [1], while an alternative choice of $X = 0.4$ has been examined in Ref. [3].

Although the standard Wilson action was used for the lattice gauge action in Ref. [1], in this paper, we extend the discussion to an improved lattice gauge action [16] in a category of actions consisting of the 1×1 plaquettes and 1×2 planar loops ("rectangles"), which are defined in the (μ, ν) -plane at a site x with the gauge link variables $U(x, \mu)$ as below.

$$W_{\mu\nu}^{1 \times 1}(x) = \frac{1}{3} \text{ReTr} [U(x, \mu)U(x + \hat{\mu}, \nu)U^\dagger(x + \hat{\nu}, \mu)U^\dagger(x, \nu)] \quad (4)$$

and

$$W_{\mu\nu}^{1 \times 2}(x) = \frac{1}{3} \text{ReTr} [U(x, \mu)U(x + \hat{\mu}, \nu)U(x + \hat{\mu} + \hat{\nu}, \nu)U^\dagger(x + 2\hat{\nu}, \mu)U^\dagger(x + \hat{\nu}, \nu)U^\dagger(x, \nu)], \quad (5)$$

TABLE I: The coefficients $C_{2,4,6}$ in the tree-level $\mathcal{O}(a^2)$, $\mathcal{O}(a^4)$ and $\mathcal{O}(a^6)$ terms of $t^2\langle E(t) \rangle$ on both the Wilson gauge configurations and the Iwasaki gauge configurations for various types of the gradient flow [9]. The values of the optimal rectangle coefficient for the $\mathcal{O}(a^4)$ -improved flows: $c_{f1,f2}^{\text{WG}}$ and $c_{f3,f4}^{\text{IG}}$ are given in the text in Sec. II C. In the table, “plaq-plus-clover” stands for an appropriate weighted average of the plaquette and clover-type of the energy density $\langle E \rangle$ with weight factors defined in Eq. (10).

types of configurations	types of gradient flow	types of $\langle E \rangle$	C_2	C_4	C_6
Wilson ($c_g = 0$)	unimproved Wilson flow ($c_f = 0$)	clover	-0.0417	-0.0020	-0.0002
	unimproved Iwasaki flow ($c_f = -0.331$)	clover	-0.7037	+0.8490	-1.5093
	unimproved Symanzik flow ($c_f = -1/12$)	clover	-0.2083	+0.0652	-0.0300
	$\mathcal{O}(a^2)$ -imp Wilson flow ($c_f = 0$)	plaq-plus-clover	0	+0.0044	+0.0014
	$\mathcal{O}(a^2)$ -imp Iwasaki flow ($c_f = -0.331$)	plaq-plus-clover	0	-0.0395	+0.1362
	$\mathcal{O}(a^2)$ -imp Symanzik flow ($c_f = -1/12$)	plaq-plus-clover	0	+0.0228	-0.0053
	$\mathcal{O}(a^4)$ -imp Wilson-like flow ($c_f = c_{f1}^{\text{WG}}$)	plaq-plus-clover	0	0	+0.0004
	$\mathcal{O}(a^4)$ -imp Iwasaki-like flow ($c_f = c_{f2}^{\text{WG}}$)	plaq-plus-clover	0	0	+0.0272
Iwasaki ($c_g = -0.331$)	unimproved Wilson flow ($c_f = 0$)	clover	-0.2623	+0.0937	-0.0480
	unimproved Iwasaki flow ($c_f = -0.331$)	clover	-0.9243	+1.2569	-2.3872
	unimproved Symanzik flow ($c_f = -1/12$)	clover	-0.4290	+0.2395	-0.1765
	$\mathcal{O}(a^2)$ -imp Wilson flow ($c_f = 0$)	plaq-plus-clover	0	+0.0408	-0.0098
	$\mathcal{O}(a^2)$ -imp Iwasaki flow ($c_f = -0.331$)	plaq-plus-clover	0	-0.2372	+0.6897
	$\mathcal{O}(a^2)$ -imp Symanzik flow ($c_f = -1/12$)	plaq-plus-clover	0	+0.0003	+0.0154
	$\mathcal{O}(a^4)$ -imp Symanzik-like flow ($c_f = c_{f3}^{\text{IG}}$)	plaq-plus-clover	0	0	+0.0156
	$\mathcal{O}(a^4)$ -imp positive rectangle flow ($c_f = c_{f4}^{\text{IG}}$)	plaq-plus-clover	0	0	-0.1033

where $\hat{\mu}(\hat{\nu})$ represents the unit vector in the direction indicated by $\mu(\nu)$.

The improved action is now given by the following form,

$$S_{\text{lat}}(U) = -\beta \left\{ (1 - 8c_{\text{rect}}) \sum_{x, \mu < \nu} W_{\mu\nu}^{1 \times 1}(x) + c_{\text{rect}} \sum_{x, \mu, \nu} W_{\mu\nu}^{1 \times 2}(x) \right\}, \quad (6)$$

which contains two parameters: the bare gauge coupling g_0 being $\beta = 6/g_0^2$ and the rectangle coefficient c_{rect} [16]. Popular choices for the value of c_{rect} yield the standard Wilson action ($c_{\text{rect}} = 0$), the tree-level Symanzik action ($c_{\text{rect}} = -1/12$ [16]) and the renormalization-group (RG) improved Iwasaki action ($c_{\text{rect}} = -0.331$ [13]), respectively.

The associated flow $V_t(x, \mu)$ of lattice gauge fields is defined by the following equation with the initial conditions $V_t(x, \mu)|_{t=0} = U(x, \mu)$

$$a^2 \dot{V}_t(x, \mu) = -g_0^2 \{ \partial_{x, \mu} S_{\text{lat}}(V_t) \} V_t(x, \mu), \quad (7)$$

where $\partial_{x, \mu}$ stands for the Lie-algebra valued derivative with respect to $V_t(x, \mu)$ and the dot-notation denotes differentiation with respect to the flow time t .

For above class of lattice gradient flows, tree-level discretization errors of $t^2\langle E(t) \rangle$ had been already studied in Refs. [9] and [10]. According the paper [9], the lattice

version of $t^2\langle E(t) \rangle$ can be expanded in a perturbative series in the bare coupling g_0 as

$$t^2\langle E(t) \rangle_{\text{lat}} = \frac{3g_0^2}{16\pi^2} \left[C(a^2/t) + \mathcal{O}(g_0^2) \right]. \quad (8)$$

The lattice spacing dependence of tree-level contribution appears in the first term, which is classified by powers of a^2/t as $C(a^2/t) = 1 + \sum_n C_{2n} \cdot a^{2n}/t^n$. The second contribution of $\mathcal{O}(g_0^2)$ represents quantum corrections beyond tree-level. Determinations of the coefficients C_{2n} depend on three building blocks: 1) a choice of the lattice gauge action for the configuration generation, 2) a choice of the lattice version of the action density, and 3) a choice of the lattice gauge action for the flow action. In Ref. [9], the $\mathcal{O}(a^{2n})$ correction terms have been determined up to C_8 for various cases of three building blocks.

For clarity, we will hereafter use a word of “ X flow”, when we adopt *the X gauge action for the flow*. For examples, we call *the Wilson flow* and *the Iwasaki flow* for choices of the Wilson and Iwasaki gauge actions for the flow.

B. A simple tree-level $\mathcal{O}(a^2)$ improvement

Following the tree-level improvement program proposed by Fodor *et al.* [9], we consider several improvements of the lattice gradient flow within choices of two

TABLE II: Simulation parameters of four ensembles generated by the Wilson gauge action. The values of the Sommer scale r_0 and lattice spacing a are taken from Ref. [17] and Ref. [1], respectively.

β (Action)	$L^3 \times T$	r_0/a	a [fm]	$\sim La$ [fm]	Statistics	$t_{0.3}/a^2$ (ours)	$t_{0.3}/a^2$ (Lüscher)
5.96 (Wilson)	$24^3 \times 48$	5.002	0.0999(4)	2.40	100	2.7968(62)	2.7854(62)
6.17 (Wilson)	$32^3 \times 64$	7.061	0.0710(3)	2.27	100	5.499(13)	5.489(14)
6.42 (Wilson)	$48^3 \times 96$	10.00	0.0498(3)	2.39	100	11.242(23)	11.241(23)
6.42 (Wilson)	$32^3 \times 32$	10.00	0.0498(3)	1.59	100	11.279(82)	N/A

different rectangle coefficients $c_{\text{rect},g}$ for the configuration generation and $c_{\text{rect},f}$ for the flow. Hereafter, we simply denote those coefficients as c_g and c_f . First of all, we describe a simple method for tree-level $\mathcal{O}(a^2)$ improvement. Let us consider C_2 coefficient of the $\mathcal{O}(a^2)$ correction term with both the plaquette- and clover-type definitions of the action density $E(t, x)$. The C_2 coefficients are given for the plaquette (C_{2p}) and clover (C_{2c}) as below [9]

$$C_{2p} = 2c_f + \frac{2}{3}c_g + \frac{1}{8}, \quad C_{2c} = 2c_f + \frac{2}{3}c_g - \frac{1}{24}. \quad (9)$$

Clearly, $C_{2p} \neq C_{2c}$ with the fixed c_g and c_f . Therefore, in order to eliminate tree-level $\mathcal{O}(a^2)$ effects, one can simply take a linear combination of two observables, which gives the corresponding C_2 coefficient as $C_{2mix} = \alpha_m C_{2p} + \beta_m C_{2c}$ [11]. Appropriate combination factors α_m and β_m can be determined under the condition of $C_{2mix} = 0$ with the normalization $\alpha_m + \beta_m = 1$ so that the coefficient of the leading term is unity and then obtained as

$$\alpha_m = 1 - 6 \left(2c_f + \frac{2}{3}c_g + \frac{1}{8} \right), \quad \beta_m = 6 \left(2c_f + \frac{2}{3}c_g + \frac{1}{8} \right), \quad (10)$$

which can eliminate C_{2mix} at any choice of c_g and c_f [12]. Therefore, the following linear combination

$$\alpha_m \langle E_{\text{plaq}}(t) \rangle + \beta_m \langle E_{\text{clover}}(t) \rangle \quad (11)$$

has no tree-level $\mathcal{O}(a^2)$ corrections [12]. This idea is quite simple as can be seen for the case of $c_g = c_f = 0$, where a weighted average of two observables, $\frac{1}{4} \langle E_{\text{plaq}}(t) \rangle + \frac{3}{4} \langle E_{\text{clover}}(t) \rangle$, would achieve tree-level $\mathcal{O}(a^2)$ improvement [26].

C. Tree-level $\mathcal{O}(a^4)$ improved gradient flows

Next, we would like to develop the aforementioned idea to achieve tree-level $\mathcal{O}(a^4)$ improvement. Taking a linear combination of $\alpha_m \langle E_{\text{plaq}}(t) \rangle + \beta_m \langle E_{\text{clover}}(t) \rangle$, the corresponding C_4 coefficient of the $\mathcal{O}(a^4)$ correction term is given by

$$C_{4mix} = \left[1 - 6 \left(x + \frac{2}{3}c_g \right) \right] C_{4p} + 6 \left(x + \frac{2}{3}c_g \right) C_{4c}, \quad (12)$$

where C_{4p} and C_{4c} represent the C_4 coefficients evaluated for the plaquette and clover-type of the energy density.

Here we introduce $x = 2c_f + 1/8$ for the sake of the following discussion. Indeed, coefficients of the higher order terms, C_4 , C_6 and C_8 , are given as polynomial functions of x in Ref. [9]. The explicit forms of both C_{4p} and C_{4c} coefficients are

$$C_{4p} = \frac{57}{32}x^2 - \frac{25}{128}x + \frac{57}{40}xz + \frac{57}{80}yz + \frac{1}{8}z + \frac{41}{2048} \quad (13)$$

and

$$C_{4c} = \frac{57}{32}x^2 - \frac{25}{128}x + \frac{57}{40}xy + \frac{57}{80}y^2 + \frac{1}{8}y + \frac{53}{2048}, \quad (14)$$

where $y = c_g - \frac{1}{4}$ and $z = c_g$ [9]. It is found that the coefficient C_4 is at most quadratic in x and the coefficient of the highest polynomial term is identical. Therefore, we are supposed to solve the following quadratic equation in terms of x so as to eliminate C_{4mix} :

$$C_{4mix} = -\frac{57}{160}x^2 - \left(\frac{1368c_g + 103}{1280} \right)x + \frac{48c_g + 41}{2048} = 0, \quad (15)$$

which leads to two kinds of the optimal coefficients c_f for any gauge configuration (c_g). Recall that these eliminate C_{2mix} and C_{4mix} coefficients simultaneously with a given rectangle coefficient c_g .

In this paper, we consider the Wilson gauge action ($c_g = 0$) and the Iwasaki gauge action ($c_g = -0.331$) for numerical simulations. For these set-ups, the optimal-flow coefficients c_f are given as

$$c_{f1}^{\text{WG}} = 0.012323, \quad c_{f2}^{\text{WG}} = -0.250261 \quad (16)$$

for the Wilson gauge configurations [12] and

$$c_{f3}^{\text{IG}} = -0.083756, \quad c_{f4}^{\text{IG}} = 0.342317 \quad (17)$$

for the Iwasaki gauge configurations. The superscripts of “WG” and “IG” found in Eq. (16) and (17) stand for the Wilson and Iwasaki gauge configurations, respectively. The second and third solutions (c_{f2}^{WG} , c_{f3}^{IG}) are close to the rectangle coefficient of the Iwasaki gauge action ($c_f = -0.331$) and the tree-level Symanzik gauge action ($c_f = -1/12$), while the first one (c_{f1}^{WG}) is very close to the zero, which corresponds to the Wilson gauge action ($c_f = 0$). Therefore we call the second and third flows as “Iwasaki-like flow” and “Symanzik-like flow”, and also the first one “Wilson-like flow”, respectively. The remaining one (c_{f4}^{IG}) are called as “positive rectangle flow” for convenience. Their $\mathcal{O}(a^2)$, $\mathcal{O}(a^4)$ and $\mathcal{O}(a^6)$ corrections terms $C_{2,4,6}$ are summarized in Table I.

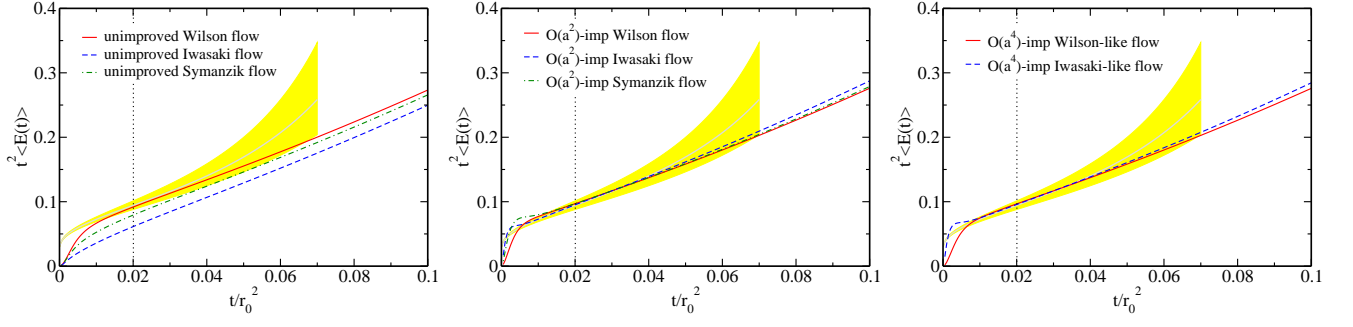


FIG. 1: (Color online) The behavior of $t^2\langle E(t) \rangle$ as functions of t/r_0^2 at $\beta = 6.17$. Three panels show results of unimproved flows (left), their $\mathcal{O}(a^2)$ -improved flows (center) and two types of $\mathcal{O}(a^4)$ improved flows (right). In each panel, the yellow shaded band corresponds to the continuum perturbative calculation [1]. A vertical dotted line on each panel marks the position of $t/r_0^2 = a^2/r_0^2$, which corresponds to the boundary of asymptotic power series expansions in terms of a^2/t .

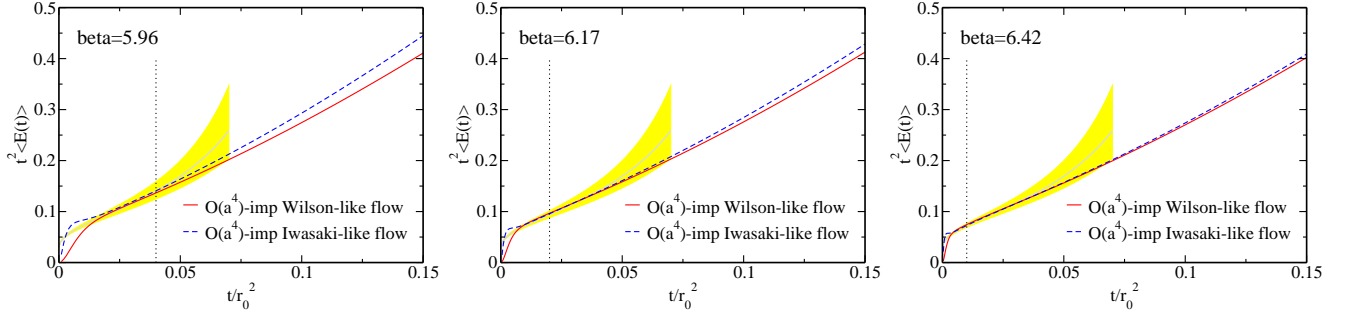


FIG. 2: (Color online) The behavior of $t^2\langle E(t) \rangle$ obtained from tree-level $\mathcal{O}(a^4)$ improved flows as functions of t/r_0^2 . Three panels show the results calculated at $\beta = 5.96$ (left), $\beta = 6.17$ (center) and $\beta = 6.42$ (right). (Same graphical conventions as in Fig. 1.)

III. NUMERICAL RESULTS

In this paper, we have performed the pure Yang-Mills lattice simulation using two different gauge actions: the standard Wilson gauge action ($c_g = 0$) and the RG improved Iwasaki gauge action ($c_g = -0.331$).

A. Results from the Wilson gauge configurations

First of all, we focus only on the results obtained from the Wilson gauge configurations and will discuss the Iwasaki gauge action results in the next subsection. The gauge ensembles in each simulation with the Wilson gauge action are separated by 200 sweeps after 2000 sweeps for thermalization. Each sweep consists of 1 heat-bath [18] combined with 4 over-relaxation [19] steps. As summarized in Table II, we generate three ensembles by the Wilson gauge action ($c_g = 0$) with fixed physical volume ($La \approx 2.4$ fm), which corresponds to the same lattice set-ups as in the original work of the Wilson flow done by Lüscher [1], and additionally generate the smaller volume ensemble ($La \approx 1.6$ fm) at $\beta = 6.42$ so as to study the finite volume effect. We have checked our code by determining a reference scale of $t_{0.3}/a^2$ from the clover-type energy density, which can be directly compared with

results of Ref. [1], as tabulated in Table II.

In the following discussion, we use five different types of the flow action for the gradient flow: Wilson, Iwasaki, Symanzik and two $\mathcal{O}(a^4)$ improved flows, and then evaluate $t^2\langle E(t) \rangle$ by means of both the plaquette- and clover-type definitions. To eliminate $\mathcal{O}(a^2)$ or $\mathcal{O}(a^4)$ corrections from the observable of $\langle E(t) \rangle$, we take the appropriate linear combinations of $\langle E_{\text{plaq}}(t) \rangle$ and $\langle E_{\text{clover}}(t) \rangle$ according to Eqs. (10) and (11). In the case of $\mathcal{O}(a^4)$ improvement, the optimal coefficients defined by the formula (16) are used. Under these variations, we classify eight different types of the gradient flow result on the Wilson gauge configurations. Their $\mathcal{O}(a^2)$, $\mathcal{O}(a^4)$ and $\mathcal{O}(a^6)$ correction terms $C_{2,4,6}$ are summarized in Table I.

In Fig. 1, we first show how our proposal of tree-level improvements works well in t -dependence of $t^2\langle E(t) \rangle$ calculated at $\beta = 6.17$. Three panels show results of unimproved flows (left), $\mathcal{O}(a^2)$ -improved flows (center) and $\mathcal{O}(a^4)$ -improved flows (right) (hereinafter, two types of tree-level improved flow are called “ $\mathcal{O}(a^2)$ -imp flow” and “ $\mathcal{O}(a^4)$ -imp flow”). A red solid (blue dashed) curve in each panel is obtained from the Wilson-type (Iwasaki-type) flows, while a green dot-dashed curve in the left and center panels is given by the Symanzik flow. The yellow shaded band in each panel represents the continuum perturbative calculation using the NLO formula of

(2) with the 4-loop $\overline{\text{MS}}$ running coupling [20] as the same prescription adopted in Ref. [1].

For the unimproved case (left panel), the Wilson flow result is closest to the continuum perturbative calculation. It is found that the larger absolute value of the rectangle coefficient c_f in the flow action further pushes the result away from the continuum perturbative calculation. This could be caused by the size of tree-level discretization errors of $t^2\langle E(t) \rangle$, which has been evaluated in Ref. [9] (as partly summarized in Table I).

Both tree-level $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ improvements indeed improve results obtained from both the Iwasaki-type and Symanzik flows significantly. Even for the Wilson-type flows, the improvements become visible in the relatively small- t regime up to $t/r_0^2 = a^2/r_0^2 \approx 0.02$, which corresponds to the boundary of asymptotic power series expansions in terms of a^2/t at $\beta = 6.17$. Furthermore, it is observed that in the range of $0.02 < t/r_0^2 < 0.05$, curves obtained from each flow almost coincide. This tendency is likely to be strong in results of the tree-level $\mathcal{O}(a^4)$ -imp flows (right panel) especially toward the smaller value of t . This indicates that the tree-level discretization errors, which may dominant in the small- t regime, are well controlled by our proposal. However, in the large- t regime ($t/r_0^2 > 0.05$), the difference between results from the Wilson-type flow ($c_f \approx 0$) and the Iwasaki-type flow ($c_f \approx -0.3$) becomes evident and also increases for a larger value of t . It is worth mentioning that at tree-level, the higher order corrections become negligible in the large- t regime due to powers of a^2/t .

What is the origin of the observed difference appeared in the larger t region? There are two major sources. One is the finite volume effect, which could be different between the results obtained from different flow actions. As mentioned before, the Yang-Mills gradient flow is a kind of diffusion equation and then the radius of diffusion becomes large as the flow time increases. Therefore, the flowed gauge fields in the larger t region are more sensitive to the boundary of the lattice. However, as we will show later, it is not the case. Another possibility is that the difference stems from some remaining discretization errors beyond the tree-level discretization effects, since non-negligible $\mathcal{O}(g^{2n}a^2)$ corrections may set in the larger t region, where the renormalized coupling g^2 becomes large.

To clarify these points, we focus on the results from two types of tree-level $\mathcal{O}(a^4)$ -imp flow. Fig. 2 displays how the observed difference between the Wilson-like flow and the Iwasaki-like flow in the large- t regime can change when the lattice spacing decreases. In Fig. 2, from the left panel to the right panel, the corresponding values of the lattice spacing in our simulations at given β are going from a coarser to a finer lattice spacing.

In Fig. 3, we also plot the difference of the values of $t^2\langle E(t) \rangle$ between $\mathcal{O}(a^4)$ -imp Wilson-like flow and $\mathcal{O}(a^4)$ -imp Iwasaki-like flow as a function of t/r_0^2 at each β . Green dot-dashed, blue dashed and red solid curves denote results at $\beta = 5.96, 6.17$ and 6.42 . This figure clearly

shows that the difference, which grows in the larger t region, becomes diminished as the lattice spacing decreases. Therefore, we confirm that the difference stems from some remaining discretization errors.

In order to determine the size of the finite volume effect, we also calculate the difference of $t^2\langle E(t) \rangle$ between two $\mathcal{O}(a^4)$ -imp flow results on the smaller lattice volumes ($32^3 \times 32$) at $\beta = 6.42$. Then, we directly compare the results obtained in two different lattice volumes ($48^3 \times 96$ and $32^3 \times 32$) as shown in Fig. 4. We confirm that there is no visible finite volume effect at least in the range of $t/r_0^2 \lesssim 0.15$.

From these observations, we conclude that the origin of the difference between two $\mathcal{O}(a^4)$ -imp flows appeared in the larger t region is caused by non-negligible $\mathcal{O}(g^{2n}a^2)$ corrections beyond the tree-level discretization effects. We then remark that the reference scale $t_{0.3}$, which is determined at around $t/r_0^2 \approx 0.1$ as originally proposed in Ref. [1], may suffer from rather large $\mathcal{O}(g^2a^2)$ errors (the order of 1% when the lattice spacing is coarse as large as $a \approx 0.1$ fm).

B. Results from the Iwasaki gauge configurations

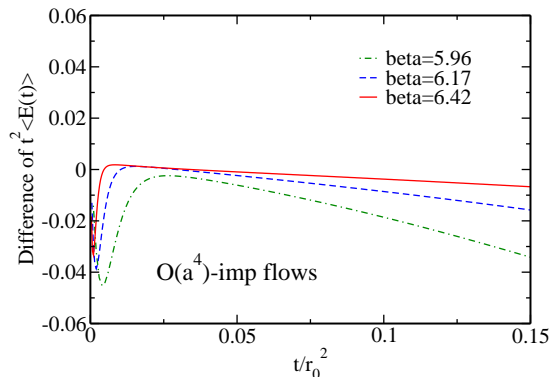


FIG. 3: (Color online) Differences of the values of $t^2\langle E(t) \rangle$ between $\mathcal{O}(a^4)$ -imp Wilson-like and $\mathcal{O}(a^4)$ -imp Iwasaki-like flows at different lattice spacings as functions of t/r_0^2 . Red solid, blue dashed and green dot-dashed curves represent results at $\beta = 5.96, 6.17$ and 6.42 .

To evaluate effectiveness of our proposal, we next study various types of tree-level improved gradient flows on the gauge configurations generated by an improved lattice gauge action including the rectangle term. We choose the Iwasaki gauge action ($c_g = -0.331$) and then generate four gauge ensembles with the similar lattice parameters (spacings a and volumes La) to the lattice set-ups for the Wilson gauge action as summarized in Table III. The Iwasaki gauge ensembles in each simulation are also separated by 200 sweeps after 2000 sweeps for thermalization as same in the cases of the Wilson gauge configurations. The smaller volume ensemble ($32^3 \times 32$) at the finer lattice spacing ($\beta = 3.10$) reserves for finite volume

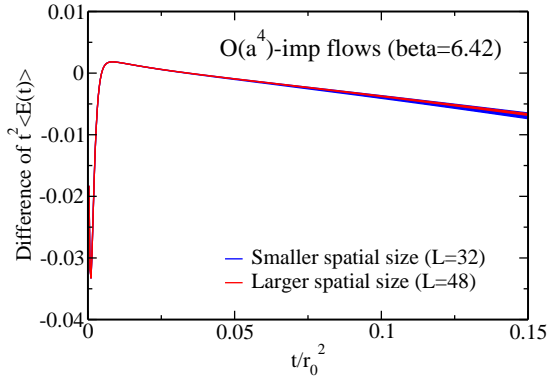


FIG. 4: (Color online) Finite volume dependence in differences of $t^2\langle E(t) \rangle$ between the $\mathcal{O}(a^4)$ -imp Wilson-like and Iwasaki-like flows as functions of t/r_0^2 . Red and blue shaded bands correspond to the results obtained in the larger ($48^3 \times 96$) and smaller ($32^3 \times 32$) volumes.

study.

We use five different types of the flow action for the gradient flow: Wilson, Iwasaki, Symanzik and two $\mathcal{O}(a^4)$ -imp flows as same in Sec. IIIB for evaluation of values of $t^2\langle E(t) \rangle$. Their $\mathcal{O}(a^2)$, $\mathcal{O}(a^4)$ and $\mathcal{O}(a^6)$ corrections terms $C_{2,4,6}$ are summarized in Table I.

Fig. 5 shows t -dependence of $t^2\langle E(t) \rangle$ calculated at $\beta = 2.80$. Three panels show results of unimproved flows (left), their $\mathcal{O}(a^2)$ -imp flows (center) and $\mathcal{O}(a^4)$ -imp Symanzik-like flow (right). For the unimproved case (left panel), the Wilson flow result is closest to the continuum perturbative calculation as same in the case of the Wilson gauge configurations. According to the size of tree-level discretization errors of $t^2\langle E(t) \rangle$ summarized in Table I, three flow results get away from the continuum perturbative result.

TABLE III: Simulation parameters of four ensembles generated by the Iwasaki gauge action. The values of the Sommer scale r_0 and lattice spacing a are taken from Ref. [21].

β (Action)	$(L^3 \times T)$	r_0/a	a [fm]	$\sim La$ [fm]	Statistics
2.60 (Iwasaki)	$24^3 \times 48$	5.078	0.0985	2.36	100
2.80 (Iwasaki)	$32^3 \times 64$	6.798	0.0736	2.36	100
3.10 (Iwasaki)	$48^3 \times 96$	10.23	0.0489	2.35	100
3.10 (Iwasaki)	$32^3 \times 32$	10.23	0.0489	1.56	100

In the center and right panels of Fig. 5, it is observed that our proposal of tree-level improvements works well as much as the case of the Wilson gauge configurations (see Fig. 1 for comparison). Here we note that we omit the result obtained from another $\mathcal{O}(a^4)$ -imp flow, namely the “positive rectangle flow”, in the right panel of Fig. 5. This is simply because the positive rectangle flow yields negative value of $\langle E(t) \rangle$ in whole positive flow time region ($t > 0$) [27]. The numerical results of the $\mathcal{O}(a^2)$ -imp Symanzik flow (center panel) and the $\mathcal{O}(a^4)$ -imp

Symanzik-like flow (right panel) mostly coincide since the C_4 coefficient of the $\mathcal{O}(a^2)$ -imp Symanzik flow is tiny as shown in Table I. Indeed, the rectangle coefficient ($c_f = -0.0838$) of the $\mathcal{O}(a^4)$ -imp Symanzik-like flow is quite close to the Symanzik action ($c_f = -1/12 \approx -0.08333$). There is neither qualitative nor quantitative difference between the results of the $\mathcal{O}(a^2)$ -imp Symanzik and $\mathcal{O}(a^4)$ -imp Symanzik-like flows. For these reasons, we hereafter focus on the $\mathcal{O}(a^2)$ -imp flows.

Let us take a closer look at the results of $\mathcal{O}(a^2)$ -imp flows (center panel). Among three types of $\mathcal{O}(a^2)$ -imp flows, the $\mathcal{O}(a^2)$ -imp Iwasaki flow becomes closest to the continuum perturbation calculation, while other flows largely overshoot the continuum counterpart in the relatively small- t regime ($0.005 \lesssim t/r_0^2 \lesssim 0.02$). This is not observed in the case of the Wilson gauge configurations, where three types of $\mathcal{O}(a^2)$ -imp flow mostly coincide near the continuum counterpart even in the small- t regime up to $t/r_0^2 \approx 0.01$. However, it should be reminded that the lattice spacing dependence of tree-level contribution is classified by powers of a^2/t as defined in Eq. 8. In the strict sense, the tree-level improvement program is supposed to be valid only in the region of $t/r_0^2 \gg (a/r_0)^2 \approx 0.02$ at $\beta = 2.80$ (Iwasaki) or $\beta = 6.17$ (Wilson).

In the large- t regime ($t/r_0^2 > 0.05$), the differences among the results from three $\mathcal{O}(a^2)$ -imp flows gradually appear and also increase for a larger value of t . As explained in Sec. III A, the origin of these differences is caused by non-negligible $\mathcal{O}(g^{2n}a^2)$ corrections beyond the tree-level discretization effects. To see this point, we show the results from three $\mathcal{O}(a^2)$ -imp flows at three different lattice spacings in Fig. 6. The left panel is for $\beta = 2.60$, the center one for $\beta = 2.80$, and the right one for $\beta = 3.10$. It is clear that the differences among three $\mathcal{O}(a^2)$ -imp flows become diminished from a coarser lattice spacing (left panel) to a finer lattice spacing (right panel).

In Fig. 7 we also plot the difference of the values of $t^2\langle E(t) \rangle$ between $\mathcal{O}(a^2)$ -imp flows as functions of t/r_0^2 . The upper (lower) panel shows the difference between $\mathcal{O}(a^2)$ -imp Wilson and $\mathcal{O}(a^2)$ -imp Iwasaki (Symanzik) flows. Green dot-dashed, blue dashed and red solid curves represent results at $\beta = 2.60, 2.80$ and 3.10 . These figures clearly show that the differences appeared in both the smaller t region ($t/r_0^2 \lesssim 0.05$) and larger t region ($t/r_0^2 > 0.05$) stem from some discretization errors. Through direct comparison of the results obtained in two different lattice volumes ($48^3 \times 96$ and $32^3 \times 32$) at the finer lattice spacing ($\beta = 3.10$), we confirm that the finite volume effects in calculations using the Iwasaki gauge configurations are also negligible in the range of $t/r_0^2 \lesssim 0.15$, as depicted in Fig. 8.

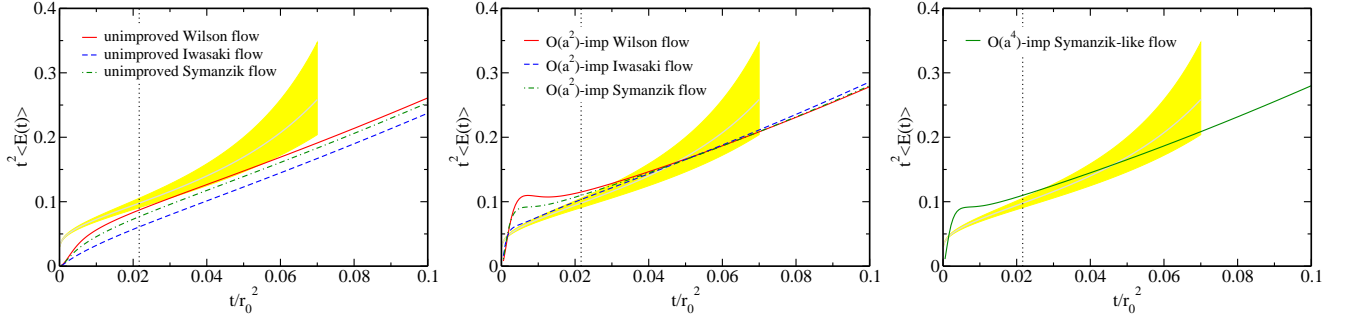


FIG. 5: (Color online) The behavior of $t^2\langle E(t) \rangle$ as functions of t/r_0^2 at $\beta = 2.80$. Three panels show results of unimproved flows (left), their $\mathcal{O}(a^2)$ -imp flows (center) and $\mathcal{O}(a^4)$ -imp Symanzik-like flow (right). (Same graphical conventions as in Fig. 1.)

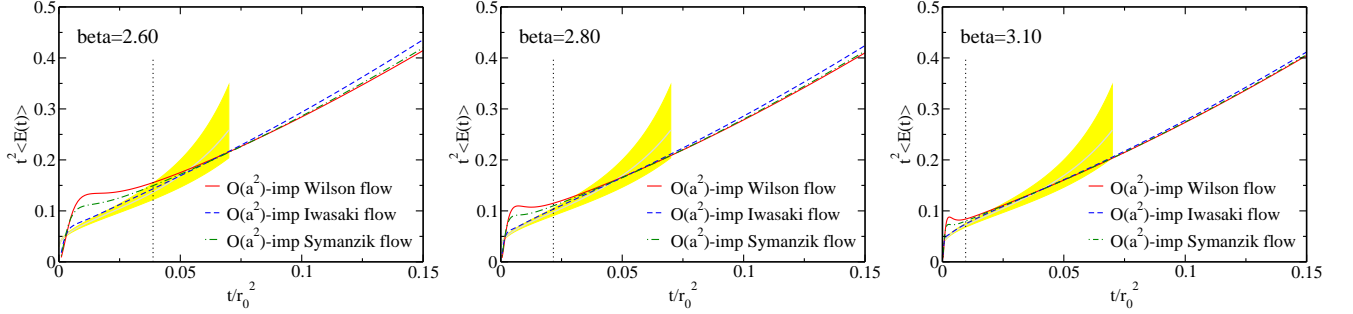


FIG. 6: (Color online) The behavior of $t^2\langle E(t) \rangle$ obtained from $\mathcal{O}(a^2)$ -imp flows as functions of t/r_0^2 . Three panels show the results calculated at $\beta = 2.60$ (left), $\beta = 2.80$ (center) and $\beta = 3.10$ (right). (Same graphical conventions as in Fig. 1.)

IV. SCALING BEHAVIOR AND CONTINUUM LIMIT OF ACTION DENSITY

In this section, we first determine the QCD parameter Λ in the $\overline{\text{MS}}$ scheme from the new reference scale t_X defined in Eq. (3) for the smaller X , where discretization uncertainties of the original (unimproved) Wilson flow become severe, in order to demonstrate the feasibility of our tree-level improvement proposal. For this purpose, we follow the analysis of Ref. [22] to evaluate the dimensionless combination of two scale parameters $\Lambda_{\overline{\text{MS}}}$ and t_X as $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$.

The parameter $\Lambda_{\overline{\text{MS}}}$ is determined through a matching between the lattice bare coupling g_0^2 to the $\overline{\text{MS}}$ running coupling g^2 with help of perturbation theory. Therefore, the choice of the smaller X , which gives the higher energy scale $\sim 1/\sqrt{t_X}$, is desirable to ensure the validity of a perturbative matching procedure.

In the previous section, we have found that the lattice discretization errors of the energy density $\langle E \rangle$ around $t/r_0^2 \approx 0.05$ are well controlled by the $\mathcal{O}(a^2)$ or $\mathcal{O}(a^4)$ -imp flows. The corresponding X is roughly 0.15, which is a fact of 2 reduced from the original choice of $X = 0.3$. In this study, we will later evaluate $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$ for $X = 0.15$ and 0.3 as typical examples.

A. Conversion to the $\overline{\text{MS}}$ scheme

As it is well known that the lattice perturbative expansions are poorly convergent, we here introduce the tadpole-improved (TI) coupling is defined by

$$g_{\text{TI}}^2(a) = \frac{g_0^2}{u_0^4} = \frac{6}{\beta u_0^4} \quad (18)$$

with

$$u_0^4 = (1 - 8c_{\text{rect}})P + 8c_{\text{rect}}R \quad (19)$$

where P and R represent the expectation values of the the path-ordered plaquette and rectangle products of link variables. The tadpole-improved coupling can *boost* the slow convergence of a power series in the lattice bare coupling.

In order to evaluate the $\Lambda_{\overline{\text{MS}}}$, let us consider a conversion from the boosted lattice scheme to the $\overline{\text{MS}}$ scheme as follows: the running coupling in the $\overline{\text{MS}}$ scheme, g^2 , is given by the following formula as a power series in the boosted coupling $g_{\text{TI}}^2(a)$, up to $\mathcal{O}(g_{\text{TI}}^4)$,

$$\frac{1}{g^2(\mu)} = \frac{1}{g_{\text{TI}}^2(a)} + 2b_0 \ln(\mu a) - t_1^{\text{TI}} + (2b_1 \ln(\mu a) - t_2^{\text{TI}}) g_{\text{TI}}^2(a) + \mathcal{O}(g_{\text{TI}}^4), \quad (20)$$

where t_1^{TI} and t_2^{TI} denote the 1-loop and 2-loop conversion variables with the first two coefficients of the β function:

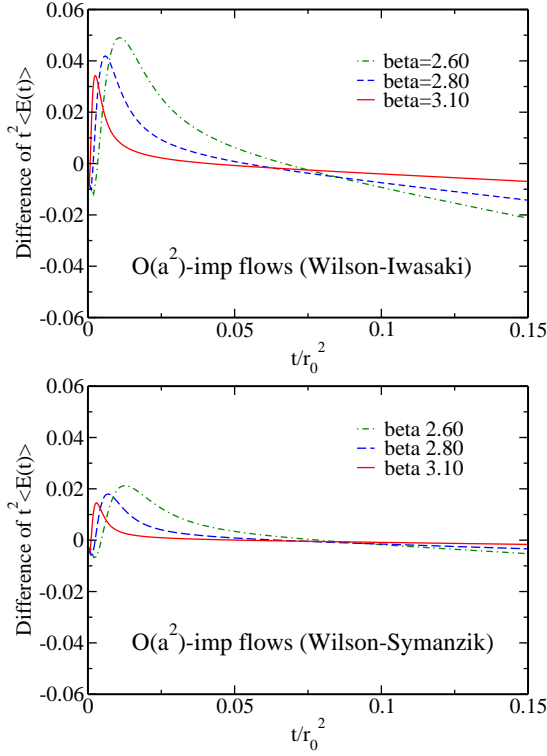


FIG. 7: (Color online) Difference of the value of $t^2\langle E(t) \rangle$ between two of three $\mathcal{O}(a^2)$ -imp flows as functions of t/r_0^2 . The left (right) panel shows the difference between $\mathcal{O}(a^2)$ -imp Wilson and $\mathcal{O}(a^2)$ -imp Iwasaki (Symanzik) flows. Green dot-dashed, blue dashed and red solid curves represent results at $\beta = 2.60, 2.80$ and 3.10 .

$b_0 = \frac{11}{(4\pi)^2}$ and $b_1 = \frac{102}{(4\pi)^4}$ being the universal coefficients in the pure Yang-Mills theory [22]. This conversion formula is fully determined by the NLO perturbation theory. However, the 2-loop conversion variable, t_2^{TI} , is not known for the case of $c_{\text{rect}} = -0.331$, since the 3-loop term of the lattice β function is not available for the RG-improved gauge action [24]. For the standard Wilson ($c_{\text{rect}} = 0$) and the Iwasaki ($c_{\text{rect}} = -0.331$) gauge actions, the currently known results of conversion variables in the tadpole improved lattice scheme are summarized in Table IV.

We next choose the renormalization scale μ to vanish the $\mathcal{O}(g_{\text{TI}}^0)$ coefficient in Eq. (20) for its rapid convergence. To achieve this, the scale μ is set to

$$\mu = \mu_* = \frac{1}{a} \exp\left(\frac{t_1^{\text{TI}}}{2b_0}\right). \quad (21)$$

This scale choice is called “method I” in Ref. [22] and reduces Eq. (20) to

$$\frac{1}{g^2(\mu_*)} = \begin{cases} \frac{1}{g_{\text{TI}}^2(a)} & (1\text{-loop}) \\ \frac{1}{g_{\text{TI}}^2(a)} + \left(\frac{b_1}{b_0} t_1^{\text{TI}} - t_2^{\text{TI}}\right) g_{\text{TI}}^2(a) & (2\text{-loop}) \end{cases} \quad (22)$$

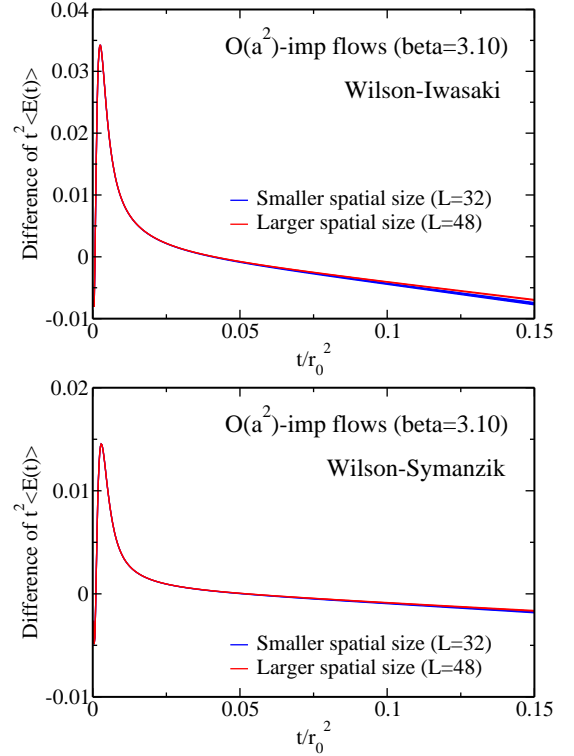


FIG. 8: (Color online) Finite volume dependence in differences of $t^2\langle E(t) \rangle$ between two of three $\mathcal{O}(a^2)$ -imp flows as functions of t/r_0^2 . The upper (lower) panel shows the difference between $\mathcal{O}(a^2)$ -imp Wilson and $\mathcal{O}(a^2)$ -imp Iwasaki (Symanzik) flows. (Same graphical conventions as in Fig. 4.)

which correspond to conversions from the boosted coupling to the $\overline{\text{MS}}$ coupling at 1-loop (first line) and 2-loop (second line) order of perturbation theory. To evaluate $g^2(\mu_*)$, we need to compute P and R numerically. We summarized our results of P and R in Table V.

Finally, we consider the standard 2-loop running coupling in the $\overline{\text{MS}}$ scheme at $\mu = \mu_*$ and thus end up with the following formula for the $\Lambda_{\overline{\text{MS}}}$ parameter:

$$\Lambda_{\overline{\text{MS}}} = \frac{1}{a} \exp\left(\frac{t_1^{\text{TI}}}{2b_0} - \frac{1}{2b_0 g^2(\mu_*)}\right) (b_0 g^2(\mu_*))^{-\frac{b_1}{2b_0^2}} \quad (23)$$

which can be determined by the value of $g^2(\mu_*)$ evaluated from either the 1-loop or 2-loop conversion formula in Eq. (22). We then calculate $\sqrt{t_X} \Lambda_{\overline{\text{MS}}}$ from Eq. (23) using the numerically computed value of t_X/a^2 . If we stress that the 1-loop (2-loop) formula in Eq. (22) is used for the conversion between two schemes, the resulting Λ parameter in the $\overline{\text{MS}}$ scheme from Eq.(23) is denoted by $\Lambda_{\overline{\text{MS}}}^{1\text{-loop}}$ ($\Lambda_{\overline{\text{MS}}}^{2\text{-loop}}$).

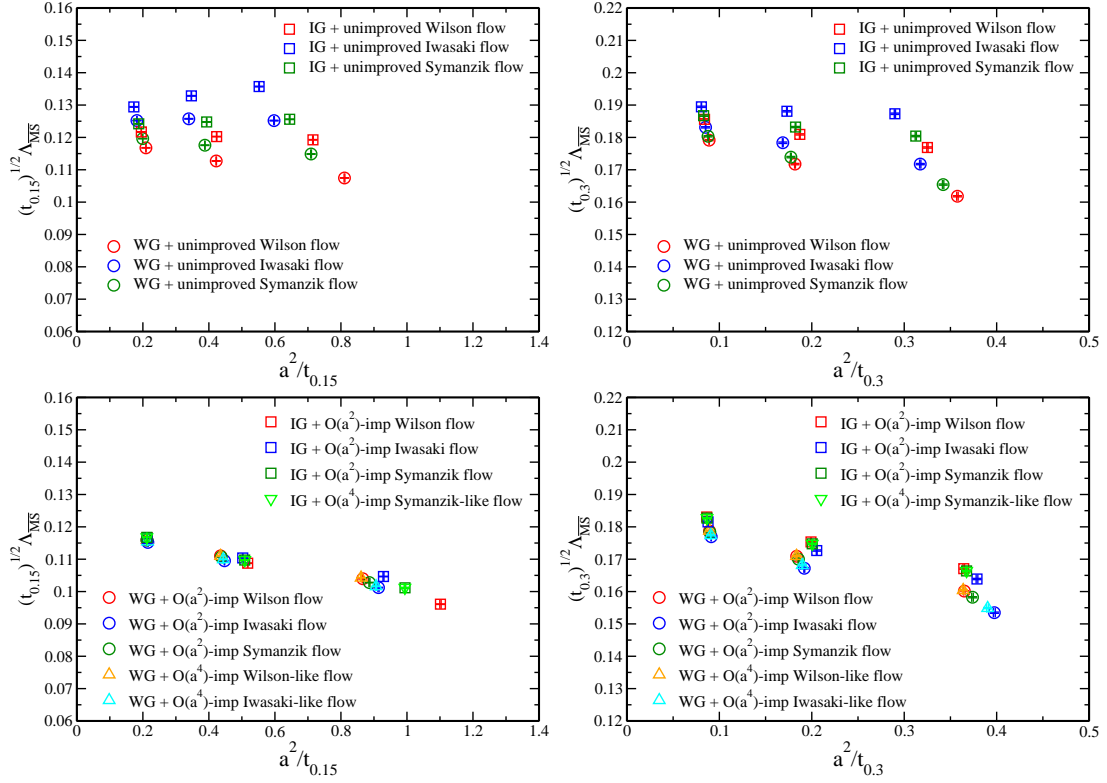


FIG. 9: (Color online) The scaling behaviors of $\sqrt{t_X} \Lambda_{\overline{MS}}$ versus a^2/t_X for choices of $X = 0.15$ (left panels) and $X = 0.3$ (right panels). Two upper panels show the results obtained by unimproved flows, while two lower panels are for the results obtained by $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ -imp flows. In legends of these figures, WG (IG) stands for the Wilson (Iwasaki) gauge action used in generating gauge configurations.

TABLE IV: The 1-loop and 2-loop conversion variables t_1^{TI} and t_2^{TI} from the boosted lattice coupling to the \overline{MS} coupling for the standard Wilson ($c_{\text{rect}} = 0$) and the RG-improved Iwasaki ($c_{\text{rect}} = -0.331$) gauge actions.

c_{rect}	t_1^{TI}	t_2^{TI}	Ref.
0	0.1348680	0.0217565	[22]
-0.331	0.1006	N/A	[23]

B. Scaling behavior of the Λ parameter

In the following discussion, we use the 1-loop formula in Eq. (22) for evaluating the value of g^2 in the \overline{MS} scheme on both the Wilson and Iwasaki gauge configurations to be treated on the same footing, since the value of t_2^{TI} is not known for the Iwasaki gauge action as mentioned earlier.

In Fig. 9, we plot the results of $\sqrt{t_X} \Lambda_{\overline{MS}}^{1\text{-loop}}$ against a^2/t_X from several combinations of two gauge actions (WG and IG) and various flows for choices of $X = 0.15$ (left panels) and $X = 0.3$ (right panels). Two upper panels show the results obtained by unimproved flows, while two lower panels are for the results obtained by

$\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ -imp flows.

For the unimproved flows (upper panels), the results for both cases of $X = 0.15$ (left) and 0.3 (right) very much depend on the choice of the gauge action and the flow. However, once any tree-level improvement is achieved, $\sqrt{t_{0.15}} \Lambda_{\overline{MS}}$ shows a nearly perfect scaling behavior as a function of a^2 regardless of types of the gauge action and the flow. For the larger X , the scaling behavior becomes less prominent. The behavior of $\sqrt{t_{0.3}} \Lambda_{\overline{MS}}$ indeed reveals a weak dependence of the choice of the gauge action, while the scaling behavior among various tree-level improved flows on the same gauge configurations

TABLE V: The expectation value of plaquette (P) and rectangle (R) values measured in this study.

β (Action)	P	R
5.96 (Wilson)	0.589 1583(32)	—
6.17 (Wilson)	0.610 8670(14)	—
6.42 (Wilson)	0.632 2170(06)	—
2.60 (Iwasaki)	0.670 6232(15)	0.452 8281(186)
2.80 (Iwasaki)	0.696 4317(61)	0.490 0710(102)
3.10 (Iwasaki)	0.727 6215(21)	0.536 2049(037)

TABLE VI: Results of the continuum value of $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$ for $X = 0.15$ and 0.3 with various types of the tree-level improved flows. The continuum extrapolation is performed by a least-square fit to all three data points or two data points given at finer lattice spacings using the linear form in terms of a^2 .

type of calculation	type of extrapolation	$\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}^{1\text{-loop}}$	$\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}^{2\text{-loop}}$	$\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}^{1\text{-loop}}$	$\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}^{2\text{-loop}}$
WG + $\mathcal{O}(a^2)$ -imp Wilson flow	3 points linear	0.1194(1)	0.1360(1)	0.1840(2)	0.2096(2)
WG + $\mathcal{O}(a^2)$ -imp Wilson flow	2 points linear	0.1206(1)	0.1368(2)	0.1861(3)	0.2111(4)
WG + $\mathcal{O}(a^4)$ -imp Wilson-like flow	3 points linear	0.1193(1)	0.1358(1)	0.1840(2)	0.2095(3)
WG + $\mathcal{O}(a^4)$ -imp Wilson-like flow	2 points linear	0.1206(1)	0.1367(2)	0.1861(3)	0.2111(4)
IG + $\mathcal{O}(a^2)$ -imp Iwasaki flow	3 points linear	0.1194(1)	N/A	0.1861(2)	N/A
IG + $\mathcal{O}(a^2)$ -imp Iwasaki flow	2 points linear	0.1209(1)	N/A	0.1882(3)	N/A

remains visible in the smaller region of a^2 . The violation of the scaling of $\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}$ as a function of a^2 might be caused by remnant $\mathcal{O}(g^{2n}a^2)$ corrections in the tree-level improved flows, which are indeed found in the larger t region, where the renormalized coupling g^2 becomes large, as discussed in Fig. 3 and Fig. 7.

Even for the case of $\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}$, where the nearly perfect scaling is achieved, there is still a slight linear dependence of a^2 . However, if one reads off the slopes of the scaling behaviors from the lower left and right panels of Figure 9, the slope for $X = 0.15$ is less steep than the case of $X = 0.3$. The origin of linear scaling in terms of a^2 is undoubtedly related to the remnant $\mathcal{O}(g^{2n}a^2)$ corrections as we will discuss in detail later.

We now consider the continuum limit of the values of $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$. Among the various flow results, we focus on the tree-level $\mathcal{O}(a^2)$ improvement flows for the specific cases ($c_g = c_f$): the $\mathcal{O}(a^2)$ -imp Wilson flow on the Wilson gauge configurations and the $\mathcal{O}(a^2)$ -imp Iwasaki flow on the Iwasaki gauge configurations.

For the continuum extrapolation, we simply adopt a linear form in terms of a^2 :

$$\sqrt{t_X}\Lambda_{\overline{\text{MS}}}(a) = (\sqrt{t_X}\Lambda_{\overline{\text{MS}}})_{\text{con}} + D_X \cdot \frac{a^2}{t_X}. \quad (24)$$

Making a least-squares fit to all three data points using Eq. (24), we get

$$\begin{aligned} (\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}^{1\text{-loop}})_{\text{con}} &= 0.1194(1) \\ (\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}^{1\text{-loop}})_{\text{con}} &= 0.1840(2) \end{aligned}$$

from WG + $\mathcal{O}(a^2)$ -imp Wilson flow and

$$\begin{aligned} (\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}^{1\text{-loop}})_{\text{con}} &= 0.1194(1) \\ (\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}^{1\text{-loop}})_{\text{con}} &= 0.1861(2) \end{aligned}$$

from IG + $\mathcal{O}(a^2)$ -imp Iwasaki flow. Even if we exclude the data point at the coarsest lattice spacing from the fit, the results are not much different as summarized in Table VI. Clearly, for the smaller X , the resulting continuum value is stable against the choice of the gauge action

and the flow. In this sense, after the tree-level improvement is achieved, the reference scale $t_{0.15}$ is much better controlled in comparison to the original one $t_{0.3}$.

Recall that we do not take into account the possible large systematic error stemming from the uncertainty of determining $\Lambda_{\overline{\text{MS}}}$. A precise determination of the continuum value of $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$ is beyond scope of the present paper. The reason why we examine the scaling behavior of $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$ is that we would like to know the scaling property purely obtained from the reference scale t_X without unknown systematic uncertainties of the lattice spacing discretization errors, arising from introduction of other observable such as the Sommer scale r_0 .

Next, we examine the uncertainty of determining $\Lambda_{\overline{\text{MS}}}$. For the Wilson gauge action, the fully NLO formula of the conversion from the boosted coupling to the $\overline{\text{MS}}$ coupling is known with the values of t_1^{TI} and t_2^{TI} as given in Table IV. Therefore, we would like to compare results of $(\sqrt{t_X}\Lambda_{\overline{\text{MS}}})_{\text{con}}$, which are determined with both the 1-loop and 2-loop conversions of the formula (22). When the 2-loop conversion is used, we get

$$\begin{aligned} (\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}^{2\text{-loop}})_{\text{con}} &= 0.1360(1) \\ (\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}^{2\text{-loop}})_{\text{con}} &= 0.2096(2) \end{aligned}$$

which indicates that the uncertainties stemming from the scheme conversion on the coupling are estimated as about 7% for the determination of $(\sqrt{t_X}\Lambda_{\overline{\text{MS}}})_{\text{con}}$.

In Fig. 10, we plot the continuum extrapolated values of $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$ for various choices of the X value as a function of X in the range of $0.1 \lesssim X \lesssim 0.4$. The results are obtained from data of WG + $\mathcal{O}(a^2)$ -imp Wilson flow. The open circle symbols represent the results from usage of the 2-loop conversion, while the open diamond symbols represent the results from usage of the 1-loop conversion. Clearly a difference of two results on each X is considerably larger than their own statistical errors, meanwhile the difference becomes more pronounced for the larger X .

Fig. 10 also includes three results from the continuum perturbation theory. At a given scale $1/\sqrt{8t_X}$, we evaluate the 4-loop $\overline{\text{MS}}$ running coupling [20] and then calculate the value of $t_X\langle E(t_X) \rangle$ using the LO, NLO and

NNLO gradient flow formula of Eq. (2). The dashed, dot-dashed and solid curves represent the LO, NLO and NNLO results. Our numerical results of $(\sqrt{t_X}\Lambda_{\overline{\text{MS}}}^{2\text{-loop}})_{\text{con}}$ from the lattice gradient flow are fairly consistent with the NNLO result from the perturbative gradient flow in the range of $0.1 \lesssim X \lesssim 0.2$.

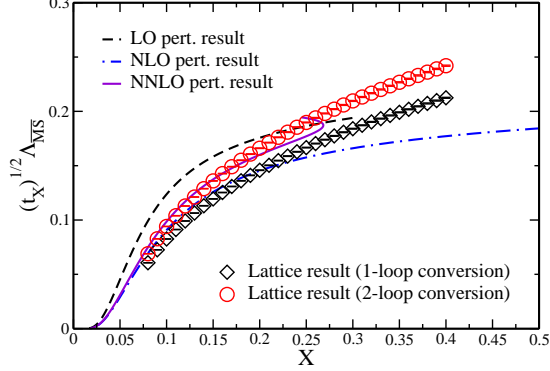


FIG. 10: (Color online) Comparison of the lattice gradient flow and the perturbative gradient flow with respect to the values of $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$ for various choices of X . For the lattice gradient flow, the continuum extrapolation is performed at fixed X . The open circle (diamond) symbols represent the lattice results obtained with the $\Lambda_{\overline{\text{MS}}}$ parameter using the 1-loop (2-loop) conversion from the lattice coupling to the $\overline{\text{MS}}$ coupling. The dashed, dot-dashed and solid curves represent the LO, NLO and NNLO results from the perturbative gradient flow with the 4-loop $\overline{\text{MS}}$ running coupling.

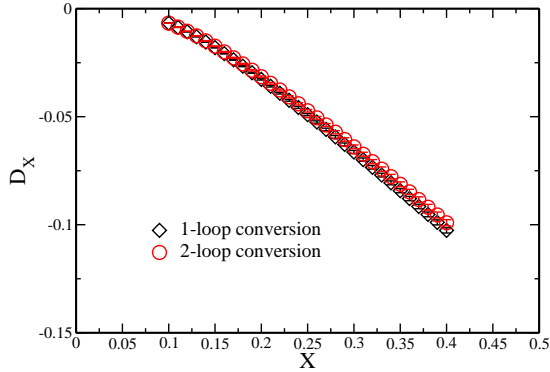


FIG. 11: (Color online) The behavior of the coefficient D_X as a function of X . D_X is coefficient of the linear term with respect to a^2/t_X in the fit form of (24).

C. Remnant $\mathcal{O}(g^{2n}a^2)$ corrections

Finally, we discuss the origin of linear scaling in terms of a^2 observed in Fig. 9. The strength of the remnant scaling violation that is approximately proportional to a^2 can be read off from the X dependence of the coefficient D_X defined in the fitting form of Eq. (24). In

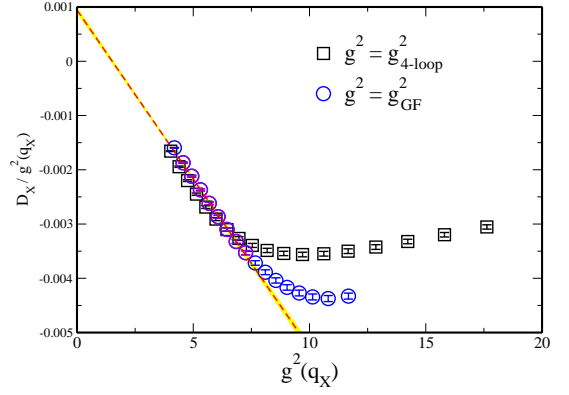


FIG. 12: (Color online) The ratio of D_X to g^2 as a function of g^2 . The value of g^2 is evaluated by two methods: the 4-loop running coupling (open squares) and the gradient flow (GF) coupling (open circles), which are defined in the text. A dashed line represents the fit result to the data points given by the GF coupling with the linear form in terms of g^2 .

Fig. 11, we show the X -dependence of the values of D_X evaluated from both results with usage of the 1-loop and 2-loop conversions of the formula (22). Although there are large uncertainties in determination of $(\sqrt{t_X}\Lambda_{\overline{\text{MS}}})_{\text{con}}$ due to the choice of the conversion formula (22), the slope coefficients D_X obtained in the case of both conversions, differences of two results for D_X are negligible especially for the smaller X region.

Therefore, in the following discussion, we use the results of D_X obtained from the 2-loop conversion with $\text{WG} + \mathcal{O}(a^2)\text{-imp}$ Wilson flow. As briefly mentioned earlier, non-zero value of D_X for given X , which is associated with the scaling violation, could be mainly caused by non-negligible $\mathcal{O}(g^{2n}a^2)$ corrections beyond the tree-level discretization effects as shown in Fig 11. It indeed seems that the value of D_X goes to zero as X decreases. If it is true, D_X might vanish because of the asymptotic freedom ($g^2 \rightarrow 0$) at high energies ($X \rightarrow 0$). Therefore, we assume that D_X is expressed by a power series in the $\overline{\text{MS}}$ running coupling g at a scale of $q_X = 1/\sqrt{8t_X}$ as below

$$D_X = \sum_n \frac{d_{2n}}{(4\pi)^n} \cdot g^{2n}(q_X) \quad (25)$$

where d_{2n} are the perturbative expansion coefficients. If the summation in above equation starts at $n = 1$ rather than $n = 0$, the origin of linear scaling is mainly associated with the remnant $\mathcal{O}(g^{2n}a^2)$ corrections, which arise beyond the tree-level.

To verify this assumption, let us plot the ratio of D_X to g^2 as a function of g^2 as shown in Fig. 12. As for the value of g^2 , we consider two types of estimation: 1) One is to use the perturbative 4-loop expression for the $\overline{\text{MS}}$ running coupling $g^2(q_X)$ with our numerical results of $(\sqrt{t_X}\Lambda_{\overline{\text{MS}}}^{2\text{-loop}})_{\text{con}}$ at given X . 2) Another is to use the gradient flow (GF) coupling g_{GF}^2 [25], which can be determined from the gradient flow formula of Eq. (2) with

a given value of $t^2\langle E(t) \rangle$. In the latter, we adopt the NNLO formula, which yields a cubic equation with respect to $g^2(q_X)$ at fixed X . We thus evaluate the ratio of D_X to g^2 using above two methods as shown in Fig. 12.

The open squared symbols represent the ratio given by the 4-loop running coupling, while the open circle symbols are evaluated using the GF coupling. When $g^2 \gtrsim 4\pi$, the cubic equation with respect to g^2 admits no feasible solution. For each case, we thus show 17 data points, which are in the range of $0.10 \leq X \leq 0.28$. In the weaker coupling regime ($g^2 \lesssim 7$), two results from different evaluations of the ratio D_X/g^2 become overlapped with each other. Their 8-9 data points ($0.1 \leq X \lesssim 0.18$) start to show the expected weak coupling scaling behavior, which is almost linear in g^2 . To ensure this point, we have carried out the linear fit on the data set given by the GF coupling, which exhibits a milder g^2 dependence, using the following expression:

$$D_X/g^2(q_X) = \frac{d_2}{4\pi} + \frac{d_4}{(4\pi)^2} \cdot g^2(q_X) \quad (26)$$

where d_2 and d_4 correspond to the coefficients for the first and second orders of g^2 in Eq. (25).

The stability of the fit results has been tested against the number of fitted data points. The best fit is drawn to fit eight data points, which are indicated by violet open circles in Fig. 12, with a reasonable value of $\chi^2/\text{d.o.f.} \approx 1.0$. We then obtain the following results:

$$\begin{aligned} d_2 &= +1.2(1) \times 10^{-2} \\ d_4 &= -9.8(2) \times 10^{-2} \end{aligned}$$

which indicate that the remnant $\mathcal{O}(g^{2n}a^2)$ corrections are reasonably small in the weak coupling regime. The fit result with 1 standard deviation is indicated by a red dashed line with a yellow shaded band in Fig. 12. From above observation on D_X , we conclude that the origin of linear scaling in terms of a^2 found in Fig 9 is related to the remnant $\mathcal{O}(g^{2n}a^2)$ corrections, which are beyond the tree-level. Although it is thus evident that the tree-level improvement program studied in this paper is not enough to eliminate the $\mathcal{O}(a^2)$ effects exactly, the remnant $\mathcal{O}(g^{2n}a^2)$ corrections can be well under control even by the simple method for the tree-level $\mathcal{O}(a^2)$ improvement.

V. SUMMARY

We have studied several types of tree-level improvement on the Yang-Mills gradient flow in order to reduce the lattice discretization errors on the expectation value of the action density $\langle E(t) \rangle$ in line with Ref. [9]. For this purpose, the rectangle term is included in both the flow and gauge actions in the minimal way. We propose a simple idea of achieving tree-level $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ improvements on $\langle E(t) \rangle$, using the linear combination of

two types of $\langle E(t) \rangle$ given by the plaquette- and clover-type definitions. For testing our proposal, numerical simulations have also been performed with both the Wilson and Iwasaki gauge configurations generated at various lattice spacings.

Our numerical results have showed that tree-level lattice discretization errors on the quantity of $t^2\langle E(t) \rangle$ are certainly controlled in the small- t regime for up to $t \gtrsim a^2$ by both tree-level $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ improved flows. On the other hand, the values of $t^2\langle E(t) \rangle$ in the large- t regime are different among the results given by different flow types leading to the same improved flow up to either $\mathcal{O}(a^2)$ or $\mathcal{O}(a^4)$ at tree-level.

In order to demonstrate the feasibility of our tree-level improvement proposal, we study the scaling behavior of the dimensionless combinations of two scale parameters, $\sqrt{t_X}\Lambda_{\overline{\text{MS}}}$, which is free from unknown systematic uncertainties regarding the lattice spacing discretization errors, arising from introduction of other observable such as the Sommer scale r_0 . For the smaller X , *e.g.* $X = 0.15$, once any tree-level improvement is achieved, $\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}$ shows a nearly perfect scaling behavior as a function of a^2 regardless of types of the gauge action and the flow. However, there is still a slight linear dependence of a^2 appeared in the cases of both $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ improved flows. On the other hand, for the larger X , *e.g.* the original choice of $X = 0.3$, the behavior of $\sqrt{t_{0.3}}\Lambda_{\overline{\text{MS}}}$ reveals a weak dependence of the choice of the gauge action, while the scaling behavior among various tree-level improved flows on the same gauge configurations remains visible especially in the smaller region of a^2 .

All aforementioned features regarding the slight scaling violation and the gauge action dependence suggest that there still remains the remnant $\mathcal{O}(g^{2n}a^2)$ corrections, which is beyond the tree level. Indeed, the origin of linear scaling in terms of a^2 found in $\sqrt{t_{0.15}}\Lambda_{\overline{\text{MS}}}$, is related to the remnant $\mathcal{O}(g^{2n}a^2)$ corrections, which can be read off from the g^2 dependence of the slope coefficient associated with the linear scaling.

Although it is evident that the tree-level improvement program studied in this paper is not enough to eliminate the $\mathcal{O}(a^2)$ effects exactly, the remnant $\mathcal{O}(g^{2n}a^2)$ corrections can be well under control even by the simple method for the tree-level $\mathcal{O}(a^2)$ improvement. Once the tree-level $\mathcal{O}(a^2)$ and $\mathcal{O}(a^4)$ improvements are achieved, the resulting energy density $\langle E(t) \rangle$ becomes very close to the continuum one in the small- t regime for up to $t \gtrsim a^2$. This offers an alternative reference scale t_X with the smaller value of X , such as $X = 0.15$. Indeed the continuum extrapolated value of $t_{0.15}$ is in excellent agreement with the perturbative gradient flow result. On the other hand, it is observed that the original reference scale $t_{0.3}$ suffers from rather large $\mathcal{O}(g^{2n}a^2)$ errors when the lattice spacing is coarse as large as $a \approx 0.1$ fm.

Acknowledgments

We would like to thank the members of the FlowQCD Collaboration (T. Hatsuda, T. Iritani, E. Itou, M. Kitazawa and H. Suzuki) for helpful suggestions and fruitful discussions. This work is in part based on Bridge++

code (<http://bridge.kek.jp/Lattice-code/>) and numerical calculations were partially carried out on supercomputer resources: SR16000 and XC40 at YITP, Kyoto University, SR16000 at KEK and LX406Re-2 at Cyberscience Center, Tohoku University.

-
- [1] M. Lüscher, JHEP **1008**, 071 (2010) [Erratum-ibid. **1403**, 092 (2014)].
 - [2] S. Borsanyi *et al.*, JHEP **1209**, 010 (2012).
 - [3] M. Asakawa, T. Hatsuda, T. Iritani, E. Itou, M. Kitazawa and H. Suzuki, arXiv:1503.06516 [hep-lat].
 - [4] A. Ramos, JHEP **1411**, 101 (2014).
 - [5] H. Suzuki, PTEP **2013**, no. 8, 083B03 (2013).
 - [6] L. Del Debbio, A. Patella and A. Rago, JHEP **1311**, 212 (2013).
 - [7] M. Asakawa *et al.* [FlowQCD Collaboration], Phys. Rev. D **90**, no. 1, 011501 (2014); Erratum: [Phys. Rev. D **92**, no. 5, 059902 (2015)].
 - [8] M. Lüscher, PoS LATTICE **2013**, 016 (2014).
 - [9] Z. Fodor *et al.*, JHEP **1409**, 018 (2014).
 - [10] A. Ramos and S. Sint, Eur. Phys. J. C **76**, no. 1, 15 (2016).
 - [11] FlowQCD Collaboration (private communication).
 - [12] N. Kamata and S. Sasaki, PoS LATTICE **2015**, 301 (2016).
 - [13] Y. Iwasaki, arXiv:1111.7054 [hep-lat].
 - [14] M. Lüscher and P. Weisz, JHEP **1102**, 051 (2011).
 - [15] R. V. Harlander and T. Neumann, JHEP **1606**, 161 (2016).
 - [16] M. Lüscher and P. Weisz, Commun. Math. Phys. **97**, 59 (1985) Erratum: [Commun. Math. Phys. **98**, 433 (1985)].
 - [17] S. Necco and R. Sommer, Nucl. Phys. B **622**, 328 (2002).
 - [18] N. Cabibbo and E. Marinari, Phys. Lett. B **119**, 387 (1982).
 - [19] M. Creutz, Phys. Rev. D **36**, 515 (1987).
 - [20] T. van Ritbergen, J. A. M. Vermaseren and S. A. Larin, Phys. Lett. B **400**, 379 (1997).
 - [21] S. Takeda *et al.*, Phys. Rev. D **70**, 074510 (2004).
 - [22] M. Göckeler, R. Horsley, A. C. Irving, D. Pleiter, P. E. L. Rakow, G. Schierholz and H. Stuben, Phys. Rev. D **73**, 014513 (2006).
 - [23] A. Ali Khan *et al.* [CP-PACS Collaboration], Phys. Rev. D **64**, 114506 (2001).
 - [24] A. Skouroupathis and H. Panagopoulos, Phys. Rev. D **76**, 114514 (2007).
 - [25] Z. Fodor, K. Holland, J. Kuti, D. Nogradi and C. H. Wong, JHEP **1211**, 007 (2012).
 - [26] Please note that this linear combination is different from tree-level $\mathcal{O}(a^2)$ “operator”, $\frac{4}{3}\langle E_{\text{plaq}}(t) \rangle - \frac{1}{3}\langle E_{\text{clover}}(t) \rangle$, proposed by Ramos and Sint [10].
 - [27] Note that although the positive rectangle flow yields positive values of $\langle E_{\text{plaq}}(t) \rangle$ and $\langle E_{\text{clover}}(t) \rangle$, the $\mathcal{O}(a^4)$ -improved combination of (11) makes $\langle E(t) \rangle$ negative.